DTIC FILE COPY

AD-A224 687

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
1990

3. REPORT TYPE AND DATES COVERED
Thesis/~~Dissertation~~

4. TITLE AND SUBTITLE
DETECTING TARGET WORDS WHILE MONITORING MULTIPLE AUDITORY INPUTS

5. FUNDING NUMBERS

6. AUTHOR(S)
JOHN M. BARKER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
AFIT Student at: Florida State University

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/CI/CIA -90-60

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFIT/CI
Wright-Ptatterson AFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release IAW AFR 190-1
Distribution Unlimited
ERNEST A. HAYGOOD, 1st Lt, USAF
Executive Officer, Civilian Institution Programs

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

DTIC
ELECTE
AUG 0 1 1990
S B D

14. SUBJECT TERMS

15. NUMBER OF PAGES
32

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
UNCLASSIFIED

18. SECURITY CLASSIFICATION OF THIS PAGE

19. SECURITY CLASSIFICATION OF ABSTRACT

20. LIMITATION OF ABSTRACT

NSN 7540-01-280-5500

90 07 31 052

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18
298-102

90-260

THE FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

DETECTING TARGET WORDS WHILE MONITORING MULTIPLE AUDITORY INPUTS

By

JOHN M. BARKER, JR.

A Thesis submitted to the
Department of Psychology in partial
fulfillment of the requirements for
the degree of Master of Science

Degree Awarded:

Summer Semester, 1990

The members of the Committee approve the thesis of John M. Barker, Jr. defended on June 1, 1990.

George Weaver
Professor Directing Thesis

Lloyd Elfner
Committee Member

Bruce Masterton
Committee member

William Oliver
Committee member

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

QUALITY INSPECTED 4

Abstract

How can humans best detect an auditory input while monitoring several inputs simultaneously? Two separate experiments were conducted, using a divided attention paradigm, to determine what factors influence target word detectability. The results from both experiments show an advantage in target detection if a person listens to one input in one ear and the other in the other ear (stereo) versus listening to both inputs in both ears (mono). Target detection was unaffected by variations in presentation rate in the range of 0.5 to 1.5 seconds. In the second experiment number of voices was examined as a factor. When both inputs were presented to each ear (mono) there was a clear advantage when listening to inputs that were recorded using two different voices (female and male) versus using only one voice (male only). However, the addition of a second voice did not improve target detection in the stereo condition. This latter finding may represent a limit on the effects of channel separation in target detection situations. Both a recency and primacy effect in terms of target detection as a function of the targets serial position in the list was found in both experiments. The results of these experiments have direct implications for practical applications, such as communication systems used by airline pilots.

It has been nearly forty years since Cherry (1953) described the problem of listening intently to one conversation while monitoring other conversations. This phenomenon, called the "cocktail party" problem, generated a tremendous amount of research attempting to understand how humans process information. Cherry pioneered the shadowing paradigm where subjects using headsets would hear two messages simultaneously, one to each ear, and verbalize one of the messages (attended channel). He found very little processing of the other message (unattended channel) and opened the gates for further research in this area by providing a clear demonstration of "selected attention". Since both ears were transmitting information about the stimulus signal received, selective processing suggested the operation of some sort of internal mechanism or process enabling one to switch attention from one ear to another.

Donald Broadbent (1958) conducted several experiments using the shadowing paradigm as well as the "dichotic listening" paradigm (subjects listen to and then attempt to recall two different inputs, one to each ear), the results of which led to the development of his theory of attention called the filter model. This model assumed that humans can process only limited amounts of information or inputs

arriving at our sensory organs. According to the model a "selective filter" limits extensive processing to the attended channel, by allowing only the attended inputs to pass. This filtering leads to minimal processing of the unattended channel, explaining the limits we have on our capacity to process simultaneously presented information.

The filter model and modifications of this theory (e.g. Treisman, 1964 & 1969) had in common an assumption that the selective aspects of attentional phenomena operate in the context of "channel" identity of information. For example, human subjects seem able to select information for processing (and reject competing information) based on physical stimulus characteristics such as ear of input (left or right), modality (auditory or visual), pitch (male or female voice), etc. (Wickens, 1984).

Subsequent theories have gone as far as eliminating the notion of a filter altogether (Ninio & Kahneman, 1974) and emphasizing, instead, time-sharing of a limited capacity central information processing system. However, while theoretical interpretations of attentional phenomena are constantly changing, there is a consensus that there are advantages to processing multiple auditory inputs through different channels as opposed to using the same channel (Van Cott & Kinkade, 1972).

After an extensive review of research on auditory information processing from 1950 to present, it was found that researchers have concentrated on the variables relating

to the level of "selective" processing of information in a listening situation. Most studies used focused attention and/or divided attention paradigms. In the focused attention paradigms subjects are instructed to ignore information coming into one ear and concentrate (focus) on the information entering the other ear. In the divided attention paradigms subjects are instructed to listen to information coming into both ears. Using both of these paradigms, it was found that manipulations of message content (Treisman, Squire, & Green, 1974), rate of presentation (Pelham, 1979), recall strategy (Bryden, 1971; Moray 1959; Treisman, 1969), and same or different voices (Shaffer & Hardwick, 1969) all effect the human listener's ability to follow one message to the exclusion of another.

One question concerning processing of simultaneous auditory inputs has not been adequately researched. Before attending to a specific input, we must decide which input to orient to. What affects our ability to select which auditory input is important to us, so we can then "tune out" the other inputs and selectively listen to the primary input. This will be referred to as the target identification phase. An example of this is when a pilot monitors two inputs over his headset. One input might be from an air traffic controller and the other his wingman (aircraft next to him in formation flying). In such situations it is equally important for the pilot to divide his attention between these two inputs and process either input when necessary.

The first question involves what factors will influence a pilot's ability to decide when and which input should be attended to (target identification). Once this question is answered previous research helps us to understand what factors will influence his ability to selectively process a particular input. Should he monitor both conversations in both ears (mono) or the controller in the left and wingman in the right ear (stereo)? Would the rate of speech used by the controller or wingman affect the pilot's ability to process the inputs? Would target identification be facilitated if the wingman used a male voice and the controller a female voice? What if the pilot needed to monitor three inputs? The answers to these questions are important both theoretically and practically.

The current studies focus on what we should selectively attend to while monitoring multiple auditory inputs. This was done by investigating identification of target words from lists of words presented simultaneously through headsets to subjects. A divided attention paradigm was used, instructing subjects to listen to all inputs equally with both ears. Previous studies that used a divided attention paradigm were concerned with attending to a particular input, not with target detection. The current study departs from most of the other research mentioned in that we are concerned with understanding factors influencing the ability to process multiple inputs and detect a pre-established target item. This is similar to what a pilot

does when he monitors several radios and responds when he hears his call sign. Previous research mainly was concerned with selectively attending to one input while excluding all other inputs, analogous to what a pilot does when actively processing one communication channel and excluding all other inputs. It is anticipated that the results will show that any means that allows inputs to be separated or distinguished (using different channels) will improve the ability of human observers to detect target words.

Experiment 1

The first experiment required subjects to detect the presence or absence of target words while monitoring, simultaneously, two different word lists through stereo headphones. These word lists were synchronized so that the subjects heard the words as "on top of each other". Both the rate of presentation of the words in a trial and the mode of presentation were manipulated in this experiment. The word lists were presented either both lists to both ears (mono) or one list to the left ear and the other to the right ear (stereo). In addition, pairs of words were presented at a rate of either every 1.5 seconds, 1.0 second, or 0.5 second. These variables were combined factorially resulting in six treatment groups.

Method

Subjects. Students from introductory psychology classes with no known hearing defects or experience monitoring multiple conversations using headphones served as

subjects in partial fulfillment of a course requirement. A total of 144 subjects (73 males & 71 females) were randomly assigned to the six different experimental treatment conditions. No attempt was made to balance the number of males and females across cells. A cell size of 24 subjects was determined based upon a desired power set at 80% (alpha=.05) and a moderate effect size (Cohen, 1969).

Apparatus and stimulus materials. Seventy lists of seven pairs of words were made up by random selection, with replacement, of AA mono-syllabic words in Thorndike-Lorge's (1944) word frequency count. No words were repeated within a list of pairs. Each list was recorded with a male voice onto a computer using an audio digital sampler (A.M.A.S. software on an AMIGA 2000 computer). The speaker did not know which words would serve as target items until after all trials were recorded. Once stored in the computer memory, each successive pair of words was synchronized and the rate of presentation was set (1.5, 1.0, or 0.5 seconds). These manipulations were accomplished by aligning and moving spectrographic representations of the audio information for each word. Ten of the lists were used for practice trials and 60 lists constituted the experimental trials for all conditions.

Next, for each rate condition each list was recorded from the computer on to a two-channel audio recorder with each word of the pair on a different channel. For the stereo presentations (member of each pair to separate ears),

stereo headphones were connected directly to the tape player. The mono presentations (each pair to both ears) was produced by inserting an electronic mixer between the headphone and the tape player. Using a between subjects design, six conditions resulted, each involving the exact same word pairs -- stereo (1.5 secs, 1.0 sec, or 0.5 sec presentation rates) and mono (1.5 secs, 1.0 sec, or 0.5 sec presentation rates).

Each list was preceded approximately 1.5 seconds before the first word pair by the trial number and a target word (e.g. "trial one, keyword dog"). The target word was present in half of the trials and was absent for the remaining trials. The position of the target for those trials in which it did occur varied randomly across serial positions 2, 4, or 6 with the constraint of an equal probability of occurrence in each position across all lists combined. In addition, for stereo conditions, the target, if in the lists, was present equally often on the left or right channel.

Procedure. Instructions were read to the subjects before the trials began and they were provided with a response sheet. They were told to listen for the presence of the target word with both ears equally (See Appendix for complete instructions). At the end of each trial they were instructed to circle "Y" if they heard the target and "N" if they did not hear the target. The stimuli were presented through stereo headphones using a Sony tape recorder. The

output volume was adjusted to a comfortable level by each subject during the practice trials.

Scoring. Since there is a 50/50 chance of a correct response, the dependent measure (detection score) was calculated by making an adjustment to percent hits. Based on signal detection theory (Wickens, 1989) the following formula was used to calculate the detection scores:

$$\text{Detection score} = 1-.25\{[FA/H] + [(1-H)/(1-FA)]\}$$
$$\text{Where; } H = \text{Hits}/30 \text{ and } FA = \text{False Alarms}/30$$

Results and Discussion

The total number of hits and false alarms were calculated for each subject and the means for hits and false alarms are presented in Table 1. Using the hits and false alarm scores for each subject, detection scores were then calculated for each target position (2, 4, & 6). Since preliminary analyses indicated no significant effects with regard to gender, all subsequent analyses were carried out on the results from males and females combined. Detection scores were then subjected to a Position x Rate x Mode of Presentation (3 x 3 x 2) mixed design analysis of variance (ANOVA). Mean and standard deviation scores are in Table 2. Post hoc pairwise comparison tests were used when appropriate.

Since the distribution of the detection scores was skewed towards values above .90 another ANOVA was conducted following an arcsine transformation of the detection scores.

Table 1

Mean Hit and False Alarm (in parenthesis) Proportions as a Function of Mode of Presentation and Rate

	Presentation Rate			
	1.5 Secs	1.0 Sec	0.5 Sec	Total
Mode				
Stereo	.74 (.09)	.75 (.10)	.69 (.11)	.73 (.10)
Mono	.62 (.18)	.60 (.14)	.59 (.13)	.60 (.15)
Total	.68 (.13)	.67 (.12)	.64 (.12)	.66 (.08)

Table 2

Mean Target Detection Scores as a Function of Serial Position, Mode of Presentation, and Rate

	Pos 2	Pos 4	Pos 6
Stereo			
1.5 Seconds	.941 (.031)[a]	.846 (.081)	.895 (.042)
1.0 Second	.939 (.030)	.845 (.069)	.904 (.039)
.5 Second	.895 (.053)	.852 (.051)	.869 (.051)
Mono			
1.5 Seconds	.886 (.047)	.730 (.134)	.790 (.090)
1.0 Second	.898 (.033)	.763 (.081)	.798 (.084)
.5 Second	.895 (.047)	.770 (.067)	.803 (.055)

[a] Standard Deviations

Since the statistical conclusions were equivalent in each case, only the analysis of the untransformed detection scores will be described and discussed.

The results of the ANOVA of detection scores are presented in Table 3. As expected a main effect of mode of presentation on target detection was found, $F(1, 138) = 84.5$, $p < .0001$, showing that target detection was better in the stereo mode of presentation compared to the mono mode. This effect was quite large in that it explained 37% of the variance in this study. Based on these results it appears that humans are better able to process information or are at least better able to detect target words if they monitor different inputs via different channels. The different channels in this situation were left and right ears.

The rate of word presentation did not show a significant effect on target detection, $F(2, 138) = 0.7$, $p = .5031$. Although it is almost certain that a much slower or faster rate of presentation could effect target detection, the presentation rates studied in this experiment spans the range of what is used in practical applications. In a real world communication setting it would be unlikely to find someone speaking faster than a word every 0.5 second or slower than a word every 1.5 seconds.

The position of the target word in a trial was associated with a significant effect on performance, $F(2, 276) = 140.6$, $p < .0001$, which explained 47% of the variance. Post hoc comparisons showed that the best

Table 3

Experiment 1 ANOVA Results

SOURCE	SS	df	MS	F	p	PVE*
Total Indep Score	1.545	143	.011	1.6		
Rate	.010	2	.005	0.7	.5031	.01
Mode	.570	1	.570	84.5	.0001	.37
Rate x Mode	.033	2	.017	2.5	.1001	.02
Pooled Residual	.932	138	.007			
Total Depend Score	1.822	288	.006	2.1		
Position (Pos)	.856	2	.428	140.6	.0000	.47
Pos x Mode	.088	2	.044	14.5	.0000	.05
Pos x Rate	.023	4	.006	2.0	.1145	.01
Pos x Rate x Mode	.008	4	.002	0.7	.6395	.00
Pooled Residual	.847	276	.003			
Total Score	3.367	431	.008			

* Percent Variance Explained

detection was for words in Position 2, then Position 6, and worst for Position 4 (*t* tests, $p < .01$ in each comparison). Typical serial position effects in a range of cognitive tasks have shown patterns similar to this. Usually there is a primacy and/or recency effect with a drop in performance for items in the middle of a list. Many of these findings involve memory tasks and it should be emphasized that the current experiment was not a memory experiment. A possible explanation of the serial position effects in the current experiment could involve differences in demands on attentional resources resulting from the processing of other items in the lists. Specifically, target words appearing in the first part of the list may be easier to detect because there is minimal interference from the processing of other words in the list. The words towards the end of the list may be easier to detect than those in the middle, but not as easy as those in the beginning, possibly because of some limited interference from words preceding the target, but with no interference from processing of subsequent items. The poorest detectability was for targets appearing in the middle of the list and this would be consistent with the possibility of interference associated with the processing of other words both before and after the occurrence of the target item.

In addition to the main effects mentioned above, there was a significant Position x Mode of Presentation interaction, $F(2, 276) = 14.45$, $p < .0001$. Figure 1

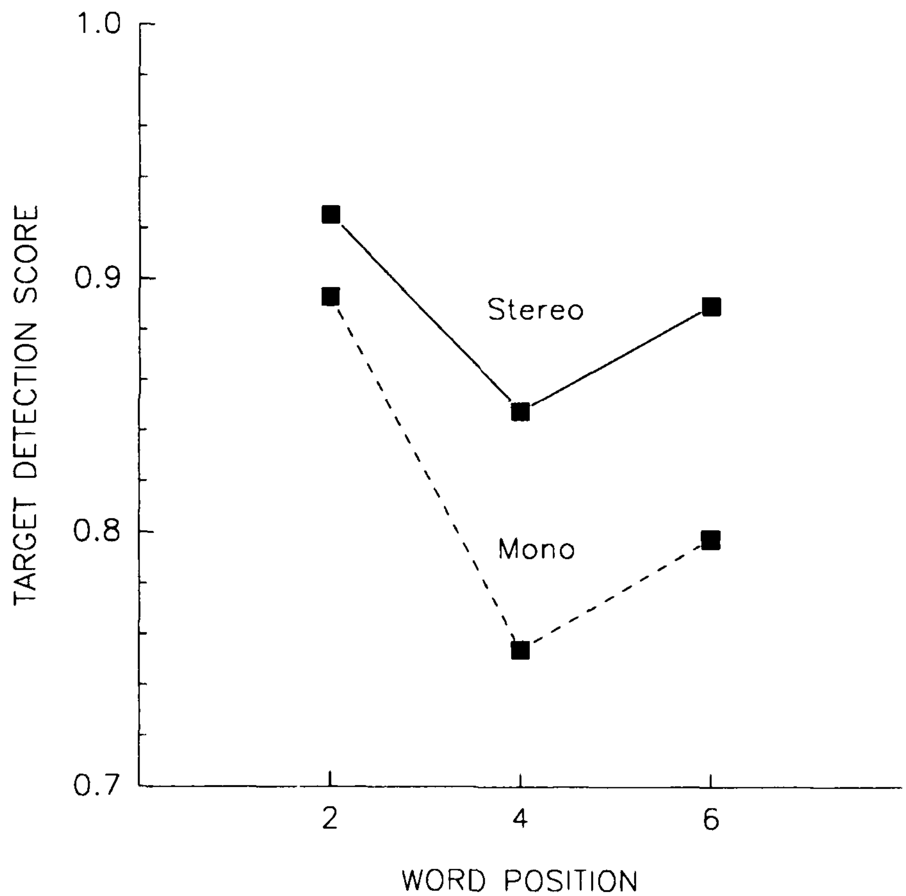

FIGURE 1. Target detection scores as a function word position and mode of presentation in Experiment 1.

illustrates this effect, which accounted for only 5% of the variance. When the target word is in Position 2 the difference in detection score is small between stereo and mono conditions (but significant), but if the word appears in either Position 4 or 6 the difference between stereo and mono is much greater. Therefore mode of presentation shows a greater effect on target detection when the actual target word appears either in the middle or towards the end of a trial versus the beginning. It is possible that this finding is the result of a ceiling effect on performance for detection of words in the earliest serial position. If so, no potential differences associated with modality (stereo vs mono) would be detected. It is possible that by adding white noise or other conditions which might make target detection more difficult, the effect of mode of presentation would be observed at all target positions.

Previous studies have shown an advantage in listening experiments for recalling verbal information when it was presented in the right ear versus the left ear. This has been labeled a right ear advantage (REA). For the current data, both left and right ear detection scores were computed and a separate analysis was used to compare these scores. The only condition where there was evidence of a REA was in the fastest rate of presentation condition (0.5 second). The results of a paired $\underline{t}$ test indicate that there were significantly more hits for target words in the right ear (11.5) versus target words in the left ear (9.2), $\underline{t}$ = 3.56,

$p < .002$. Prior studies in which the REA is obtained typically involve presentation rates of the same magnitude. The failure to show REA in the slower conditions, could be interpreted as an indication of ample time to process all information coming to both ears. But it does appear in the fast condition, presumably because there is not enough time to adequately process all inputs and therefore those arriving at the right ear will have an advantage over those arriving in the left ear.

Experiment 2

The results from Experiment 1 indicate that the mode of presentation of auditory stimuli can affect the ability of subjects to detect target words. By increasing the number of channels for inputs as in the stereo condition, performance improves. How else might we increase the number of channels and will the effects of adding channels be additive (in terms of target detection)?

To answer these questions a second experiment was conducted. Since the manipulation of the rate of presentation did not show any significant results in the first experiment it was held constant at 1.0 second for this study. Mode of presentation was manipulated once again with both stereo and mono conditions. A new factor called "voice" was added to study the effects of adding another channel, this time with stimuli differing along the pitch dimension. In the single voice condition all words were recorded using the same male voice and in the dual voice

condition, one word in each pair was recorded in a male voice and the other in a female voice. The voice conditions were balanced across trials, so that half the time the target word was in a female voice and the other half were in a male voice. Thus four conditions resulted using the exact same word pairs -- stereo (single or dual voice) and mono (single or dual voice). The expected results should show us the same effects as in Experiment 1 with regard to comparison of stereo and mono presentation mode and, in addition, the dual voice condition would be expected to increase target detection scores in both the stereo and mono conditions.

Method

Subjects. Subjects were selected in the same fashion as in Experiment 1. Based again on a power analysis a total of 128 subjects (64 males & 64 females) was used with 32 in each condition (Cohen, 1969). No attempt was made to balance the number of males and females in each condition.

Stimulus materials. The exact same word lists that were prepared for Experiment 1 were employed for the single voice condition of this experiment. To create the dual voice condition, half of the stimulus word lists were recorded into the computer using a female voice. The digital sampler was used to combine these words with the remaining word lists spoken in a male voice to produce list pairs for the dual voice condition. These list pairs were then recorded onto a two-track tape player, each voice on a

separate track. The voices were counterbalanced across left and right tracks.

Procedure. Stimuli was once again presented through stereo headsets using a Sony tape recorder and subjects were instructed to listen for the presence of the target word with both ears equally. At the end of each trial they were instructed to circle "Y" on their response sheet if they heard the target and "N" if they did not hear the target (See Appendix). Subjects used the same response sheet that was used in the first experiment.

Scoring. The same procedure as in Experiment 1 was used to provide a correction for guessing.

Results and Discussion

The total hits and false alarms were calculated for each subject (see Table 4). This data was then used to calculate detection scores by subject for each word position. Since preliminary analyses indicated no significant effects associated with gender, the subsequent analyses were carried out on the results of both sexes combined. Detection scores were subjected to a Position x Voice x Mode of presentation (3 x 2 x 2) mixed design ANOVA. Means and standard deviations of the detection scores appear in Table 5. Post hoc pairwise comparison tests were used when appropriate.

As in Experiment 1 a separate ANOVA was conducted using an arcsine transformation of the detection scores. The only statistical inference which differed for the transformed and

Table 4

Mean Hit and False Alarm (in parenthesis) Proportions as a Function of Mode of Presentation and Voice

	Voice		
	Dual Voice	Single Voice	Total
Mode			
Stereo	.76 (.10)	.79 (.09)	.78 (.10)
Mono	.77 (.17)	.59 (.16)	.68 (.17)
Total	.77 (.14)	.69 (.13)	.73 (.13)

Table 5

Mean Target Detection Scores as a Function of Serial Position, Mode of Presentation, and Voice

	Position 2	Position 4	Position 6
Stereo			
Single Voice	.942 (.038)[a]	.882 (.048)	.905 (.055)
Dual Voice	.927 (.047)	.867 (.051)	.910 (.053)
Mono			
Single Voice	.870 (.049)	.754 (.096)	.788 (.063)
Dual Voice	.915 (.040)	.838 (.054)	.869 (.070)

[a] Standard Deviations

untransformed scores was associated with the interaction of Voice and Position. For ease in interpretation, the analysis of untransformed scores will serve as the basis for the discussion which follows, except for the single case where a difference in statistical outcomes was obtained.

The results of the ANOVA using the untransformed target detection scores are presented in Table 6. As in the first experiment a main effect of mode of presentation on target detection was found, $\underline{F}(1, 124) = 780$, $\underline{p} < .0001$, showing improved target detection in the stereo condition versus the mono condition. This accounted for 32% of the variance in the ANOVA and validates the notion that using ear as a channel for input increases our ability to process auditory information.

Also as in Experiment 1 there was a significant effect of serial position on target detection, $\underline{F}(2, 248)$, $\underline{p} <.0001$, which explained almost 40% of the variance. Post hoc comparisons revealed that once again subjects were better able to detect target words when they were in Position 2, then Position 6, and were least likely to detect the words that were in Position 4.

As was expected, the effect of voice was also significant, $\underline{F}(1, 124) = 166$, $p = .0001$, explaining 7% of the variance. This finding adds support to the idea that pitch defines another channel of input, which can effect our ability to detect target words in an auditory task. It should be noted that this experiment utilized extreme

Table 6

Experiment 2 ANOVA Results

SOURCE	SS	df	MS	F	p	PVE*
Total Indep Score	1.330	127	.011	19.3		
Voice	.090	1	.090	166.3	.0001	.07
Mode	.422	1	.422	779.9	.0000	.32
Voice x Mode	.147	1	.147	271.7	.0000	.11
Pooled Residual	.671	124	.001			
Total Depend Score	1.003	256	.004			
Position (Pos)	.396	2	.198	88.2	.0000	.39
Pos x Mode	.030	2	.015	6.7	.0017	.03
Pos x Voice	.013	2	.006	2.9	.0636	.01
Pos x Voice x Mod	.007	2	.003	1.6	.1930	.01
Pooled Residual	.557	248	.002			
Total Score	2.330	383	.006			

* Percent Variance Explained

variations in pitch (female to male) and the results may be less dramatic if one used only slight variations in pitch, such as two different male voices.

Probably more interesting than the main effects were two significant interaction effects. As in the first experiment there was a Position x Mode of Presentation effect, $\underline{F}(2, 248) = 6.7$, $\underline{p} = .0017$. Figure 2 illustrates this effect, showing that mode of presentation has a greater effect on target detection when the actual target word appears either in the middle or toward the end of a trial versus the beginning. These findings replicate what was found in Experiment 1, further emphasizing that serial position has an effect on how information is processed. The possibility that this interaction is due to a potential ceiling effect remains, as explained in the discussion of Experiment 1. Consistent with the view that a ceiling effect may be present in the untransformed data was the finding that, following the arcsine transformation, this interaction failed to achieve statistical significance, $\underline{F}(2, 248) = 2.9$, $\underline{p} =.0544$. Since the transformation helps to minimize the impact of a ceiling effect by making the distribution of the scores more normal, a replication of this experiment with added noise (to reduce a ceiling effect) would be expected to mimic the results of the transformed scores in this experiment for the Position by Mode of Presentation interaction.

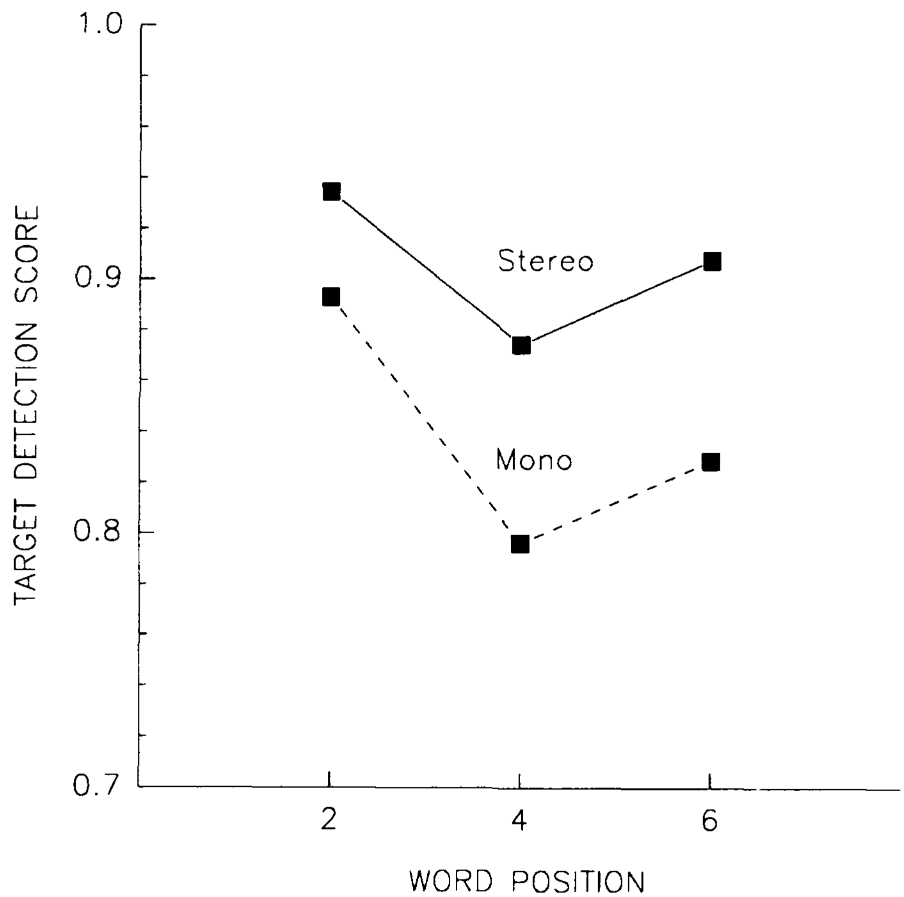

FIGURE 2. Target detection scores as a function of word position and mode of presentation in Experiment 2.

The addition of the voice factor resulted in another significant interaction effect. Figure 3 shows the Voice x Mode of Presentation effect, $F(2, 248) = 272$, $p < .0001$, which explains 11% of the variance. Post hoc comparisons show that mode of presentation had no effect on target detection in the dual voice condition, but was a reliable source of variation in the single voice condition ($p < .01$). Assuming that we have not reached a ceiling effect for target detection scores, these results would indicate that there is a limit to performance improvement that can be achieved by increasing the number of different channels for auditory inputs. Specifically, these data suggest that in the stereo presentation mode, the addition of a voice channel has no positive effect and that a limit had been reached on a person's cognitive processing capability. One way to verify that the data reflect such a processing limit rather than a ceiling effect would be to repeat the experiment, adding white noise to all conditions to lower absolute performance levels, and determine if this interaction still holds.

A separate analysis was conducted on left and right ear detection scores to determine if there was any evidence of a REA in this experiment. The results showed that there was no REA ($p = .432$) in any of the conditions for Experiment 2. This is understandable since the words were presented at a 1.0 second rate, a rate which showed no significant REA effect in the first experiment.

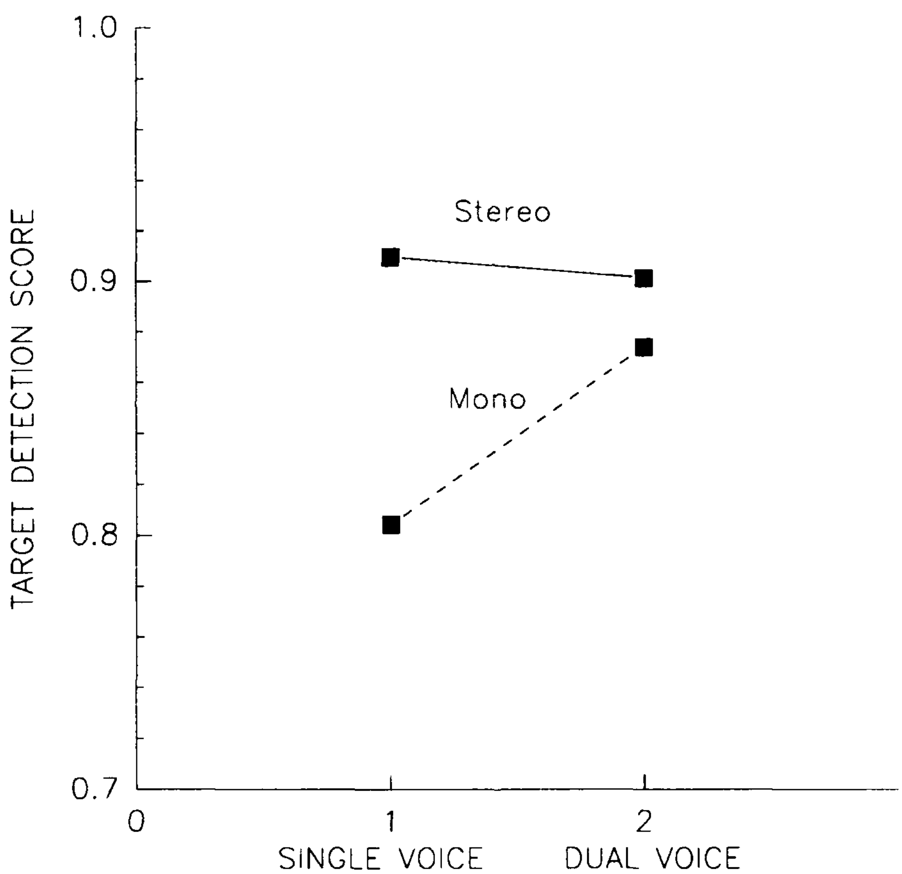

FIGURE 3. Target detection scores as a function of voice and mode of presentation in Experiment 2.

General discussion

The results of both experiments have shown that there is an advantage in detecting a target word when listening to two inputs, one to each ear (stereo) versus both inputs to both ears (mono). This supports the theory that by adding channels we can increase our cognitive processing capacities for target detection. The major difference in these experiments from previous studies was the dependant measure studied. These experiments were concerned with detecting a target word, something that might indicate that a particular conversation should then be attended to. Most other studies were concerned with actually following one conversation while excluding all other inputs. It is important to emphasize that these differences were not only statistically significant but relatively large effects. This translates to (if we can make the leap) a practical advantage in the real world. Clearly, pilots whose equipment requires them to listen with both ears to all inputs combined, similar to the mono condition, should change to a stereo method (separate channels to each ear) for monitoring communications. But pilots who use only one ear to monitor outside communications (headphone with only one ear-piece) will not benefit from these changes. This leads us to the findings of the second experiment.

The information gained from Experiment 2 shows that pitch as "channels" may be similarly effective as using different ears (sound localization) as "channels". What we

would have hoped to find was an additive effect of the voice condition and the mode of presentation condition. These effects were not additive as the voice manipulation had no effect in the stereo condition. This may have been caused by an experimental design artifact. It is possible that in the nature of the stimulus materials or experimental conditions gave rise to a ceiling effect on target detection scores, thereby restricting any advantages gained from adding another channel in the stereo condition. Alternatively, one channel dimension may provide all the advantage that can be obtained due to our limited cognitive processing capacity, thus making additional dimensions irrelevant. A reasonable future experiment would involve adding background noise to the conditions in hope of determining whether the results of Experiment 2 were the reflection of a performance ceiling or an indication of "diminishing returns" as additional channels are added. One could argue that in the real world background noise would be present, so adding it to the experiment should not affect its external validity.

These experiments showed that rate of speech did not affect target detection. Once again we must make the point that the rates used were of a limited range, but mimic what is used in practical applications. The position of a target word did have an effect. This effect showed somewhat of a classical serial position effect, indicating that if you want to improve target detectability, place the target in

either the beginning or end of a list of words. This finding supports the current practice (at least in an aviation setting) of the important target detection information being stated in the beginning of a transmission (e.g. aircraft call sign).

The interaction between serial position and mode of presentation was found in both experiments, but must be interpreted with caution. Even though the second experiment replicates the finding of the first experiment, it is possible that the design of the experiments were the actual factor causing the interaction. A simple extension of these experiments with added noise would aid in verifying whether the effects of channel differentiation are truly different as a function of the position of a target in a list of words.

Overall there is evidence that a human's ability to detect target words can be increased by varying auditory inputs along different channels. There is also support for the idea that this improvement may be limited by our cognitive capacity to process information. As mentioned, further manipulations and refinements of the current experiments would be expected to shed more light on the question of how humans can most efficiently detect target words when listening to multiple, simultaneous auditory inputs.

References

Broadbent, D. E. (1958). Perception and communication. London: Pergamon Press.

Bryden, M. P. (1971). Attentional strategies and short-term memory in dichotic listening. Cognitive Psychology, 2, 99-116.

Bryden, M. P. & Murray, J. E. (1985). Toward a model of dichotic listening performance. Brain and Cognition, 4, 241-257.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. Journal of the Acoustical Society of America, 25, 975-979.

Cohen, J. (1969). Statistical power analysis for the behavioral sciences. New York: Academic press, inc.

Moray, N. (1959). Attention in dichotic listening: Effective cues and the influence of instructions. Quarterly Journal of Experimental Psychology, 11, 56-60.

Ninio, A. & Kahneman, D. (1974). Reaction time in focused and in divided attention. Journal of Experimental Psychology, 103 (3), 394-399.

Pelham, W. E. (1979). Selective attention deficits in poor readers? Dichotic listening, speeded classification, and auditory and visual central

and incidental learning tasks. _Child-Development_, _50_ (4), 1050-1061.

Rawasi, F. & Azodi, V. (1988). _A. M. A. S._ [computer program]. Future Vision.

Shaffer, L. H. & Hardwick, J. (1969). Monitoring simultaneous auditory messages. _Perception & Psychophysics_, _6_ (6B),401-404.

Treisman, A. M. (1964). Selective attention in man. _British Medical Bulletin_, _20_, 12-16.

Treisman, A. M. (1969). Strategies and models of selective attention. _Psychological Review_, _76_ (3), 282-299.

Treisman, A. M., Squire, R., & Green, J. (1974). Semantic processing in dichotic listening? A replication. _Memory & Cognition_, _2_ (4), 641-646.

Van Cott, H. P. & Kinkade, R. G. (1972). _Human engineering guide to equipment design_. Washington, DC: U.S. Government Printing Office.

Wickens, C. D. (1984). _Engineering psychology and human performance_. Columbus, OH: Charles E. Merrill Publishing Company.

Appendix

Instructions to subjects

Your task will be to listen to a set of recorded words and identify if a keyword is present. An example of what you will hear will be "Trial 1 keyword dog.. cat ball tall tip coat etc." If you hear dog again circle Y if not circle N. Then you will hear "Trial 2 keyword fish.. did flash pet etc..." This will continue until you reach the last trial and at that time you should remove your headset and return to this room. The words may sound a little confusing since there are actually two lists recorded together. Please listen intently and equally with both ears. Answer as best as you can. If the tape stops or you can not hear anything in either on of your ears let me know. Any questions?

References

Broadbent, D. E. (1958). Perception and communication. London: Pergamon Press.

Bryden, M. P. (1971). Attentional strategies and short-term memory in dichotic listening. Cognitive Psychology, 2, 99-116.

Bryden, M. P. & Murray, J. E. (1985). Toward a model of dichotic listening performance. Brain and Cognition, 4, 241-257.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. Journal of the Acoustical Society of America, 25, 975-979.

Cohen, J. (1969). Statistical power analysis for the behavioral sciences. New York: Academic press, inc.

Moray, N. (1959). Attention in dichotic listening: Effective cues and the influence of instructions. Quarterly Journal of Experimental Psychology, 11, 56-60.

Ninio, A. & Kahneman, D. (1974). Reaction time in focused and in divided attention. Journal of Experimental Psychology, 103 (3), 394-399.

Pelham, W. E. (1979). Selective attention deficits in poor readers? Dichotic listening, speeded classification, and auditory and visual central

and incidental learning tasks. Child-Development, 50 (4), 1050-1061.

Rawasi, F. & Azodi, V. (1988). A. M. A. S. [computer program]. Future Vision.

Shaffer, L. H. & Hardwick, J. (1969). Monitoring simultaneous auditory messages. Perception & Psychophysics, 6 (6B),401-404.

Treisman, A. M. (1964). Selective attention in man. British Medical Bulletin, 20, 12-16.

Treisman, A. M. (1969). Strategies and models of selective attention. Psychological Review, 76 (3), 282-299.

Treisman, A. M., Squire, R., & Green, J. (1974). Semantic processing in dichotic listening? A replication. Memory & Cognition, 2 (4), 641-646.

Van Cott, H. P. & Kinkade, R. G. (1972). Human engineering guide to equipment design. Washington, DC: U.S. Government Printing Office.

Wickens, C. D. (1984). Engineering psychology and human performance. Columbus, OH: Charles E. Merrill Publishing Company.